基于矢量量化的高效随机物理层密钥提取方案

李 鑫,李兴华,杨 丹,马建峰
(西安电子科技大学计算机学院,陕西西安 710071)

摘 要: 针对现有的物理层密钥生成方法中存在的生成速率偏低、误码率高等问题,借助无线信号接收信号强度RSS,提出了基于矢量量化的高效随机物理层密钥提取方案(HRVQ).该方案通过不一致性去除减少通信双方不一致的信道特征值,利用矢量量化将信道信息转化为0、1比特流,并通过模糊提取器进行纠错和随机性增强处理.实验表明:该方案在密钥生成速率方面达到了284%的比特生成率,并且在实现了零误码率的同时保证了生成密钥的随机性.

关键词: 物理层安全; 信息论安全; 密钥提取; 矢量量化; 模糊提取
中图分类号: TP393.08 **文献标识码:** A **文章编号:** 0372-2112 (2015)02-0275-07
电子学报URL: http://www.ejournal.org.cn **DOI:** 10.3969/j.issn.0372-2112.2016.02.005

HRVQ:A High-Speed Random Key Extraction Scheme Based on Vector Quantization

LI Xin, LI Xing-hua, YANG Dan, MA Jian-feng
(*School of Computer Science and Technology, Xidian University, Xi'an, Shaanxi* 710071, *China*)

Abstract: To solve the dilemma that a high bit generation rate and a low probability of bit disagreement can't be achieved simultaneously in current physical layer key extraction schemes, a high-speed random key extraction scheme based on vector quantization (HRVQ) which just employs the RSS of Wi-Fi signal is proposed. First, the inconsistent RSS measurements between the two communication parties were reduced by inconsistency removal to achieve low bit disagreement rate. Then, the vector quantization was introduced to accelerate the extraction of bit streams from channel information and achieve high bit generation rate. Last, fuzzy extractor was employed to achieve adequate randomness and zero bit disagreement rate. The result of experiments indicates that our scheme outperforming the best scheme of the state-of-the-art can reach a 284% bit generation rate with a zero bit disagreement rate and a guaranteed randomness.

Key words: physical layer security; information theory security; key extraction; vector quantization; fuzzy extractor

1 引言

无线网络已经融入了人们的生活之中.然而,无线网络的广播特性使其面临比有线网络更大的安全问题,而且现有的无线加密体系也存在许多不足之处:(1)现有的802.11协议存在诸多问题[1],例如没有为管理帧和控制帧提供保护;(2)现有安全系统的理论基础是数学上的困难问题,随着计算技术的发展,攻击者的计算能力大幅提升,这些基于计算复杂性的密码技术的安全性就会降低;(3)传统的安全机制需要密钥管理、分发、更新和维护,随着节点个数的增加,所需密钥个数呈指数级增加,密钥分发和更新的工作量会非常大.因此,引入物理层安全[2]来补充或增强无线网络安全的方案受到了越来越多的关注.

物理层密钥是基于信息论安全[3]的,它能够降低密钥分配和更新的难度,实现跨层安全,为上层密钥生成提供素材,提高现有的密钥体系的安全性.例如,可以使用物理层密钥增强现有的802.11i的安全性[4].物理层信道信息包括到达角度[5]、相位[6]、接收信号强度(RSS)[7~10]、信道脉冲响应(CIR)[11]、信号包络[12,13]和电平交叉点等.RSS是提取密钥最常用的无线信道特征,因为它在现有的无线基础设施中是已经存在的,很

收稿日期:2014-07-09;修回日期:2015-02-28;责任编辑:蓝红杰
基金项目:国家自然科学基金(No. U1135002, No. 61372075, No. 61202389, No. 61100230, No. 61309016);国家密码发展基金(No. MMJJ201201004);地理信息国家重点实验室开放课题(No. SKLGIE2013-M-4-1)

容易测量得到.

Shannon[3]在 1949 年提出了信息理论安全并且证明了完全保密密钥是存在的,典型的例子就是“一次一密”. Hershey[6]等人提出在无线通信中通信双方的无线信道特征具有随机性、时空唯一性和互易性,可以作为随机信号源.

Suman Jana[8]等提出了自适应的密钥生成方案 ASBG,该方案使用自适应的量化方法,将 RSS 划分到多个区间. Neal Patwari[9,10]等先后提出了 HRUBE 方案和 ARUBE 方案,采用了 KLT 变换、多区间量化、格雷码纠错、rank 排序等方法. Hongbo Liu[14]等在所提出的方案中,以多区间量化和格雷码纠错为基础,将信号衰减趋势纳入了量化范围. 为了解决移动设备不在各自的通信范围之内的问题,Hongbo Liu[15]等采用接力节点提出解决方案. Lihua Dou[16]等从多用户角度提出了解决方案.

上述方案中的量化方法大致可分为单比特量化和多比特量化两种,单比特量化[7,12]误码率低,密钥生成速率也较低;多比特量化[8,9]可以提高密钥生成速率,同时也提高了误码率. 因此,密钥生成速率和密钥误码率之间存在矛盾.

本文提出了一种高效的物理层密钥提取方法,包括不一致性去除、矢量量化、模糊提取等步骤. 为了能够充分利用现有的硬件,使得所设计的方案具有普适性,本文借助 RSS 来提取密钥.

2 攻击者模型

在攻击者模型中,我们假定 Eve 可以窃听 Alice 和 Bob 之间的所有通信信息. 我们假设 Eve 知道密钥提取算法以及算法中相应的参数值,并且可以在测量过程中到处移动,也可以移动 Alice 和 Bob 之间的物体来改变信号的衰减,这种移动既不能增大相关时间也不能影响 Alice 和 Bob 之间的信号相关性. 但是,Eve 不能处在过于接近 Alice 或 Bob 的位置,须至少在 $\lambda/2$ 距离之外(例如:在 2.4GHz 的频率下,$\lambda/2 = 6.25$cm),才能使 Eve 测量到的信号衰落独立于 Alice 及 Bob 的信号衰落. 在本文中,我们假定攻击者与合法节点的距离在几个波长之外,在移动的环境中与攻击者保持 20cm 左右的距离并不难实现. 同时,我们假设在测量过程中 Eve 不能发出干扰信号,不能发起中间人攻击. 因此,本文提出的方案同已有方案一样,只能够抵挡被动攻击.

3 方案概述

物理层密钥有以下三点要求:(1)相同的序列:加密双方生成的密钥须完全一致;(2)合适的长度:满足对称加密算法中常用的密钥长度(一般为 128bit 至 512bit);(3)统计学随机:比特之间相互独立. 这就要求密钥生成方案满足误码率低、生成速率高、密钥随机等要求.

针对这些要求本文提出一种新的无线物理层密钥生成方法. 假定双方已经获得了相关性很高的 RSS,我们的方案框架如图 1 所示,共分为三个步骤:(1)不一致性去除,去除由信道半双工或周围噪声引起的不一致信道信息;(2)矢量量化,将 RSS 根据平均线划分成两个区间,最终把不连续的 RSS 值转化为 0、1 比特流;(3)模糊提取,对生成比特流中不一致的位进行纠错并提取出随机的比特串.

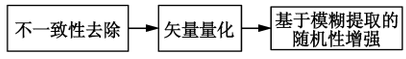


图1 HRVQ方案框架

3.1 不一致性去除

由于物理层信道的半双工特性,Alice 和 Bob 测量的微小时延会导致他们所处的相对位置发生改变,引起通信双方信道信息的不一致. 同时在 Alice 和 Bob 周围存在着一些噪声,也会在所测量的信道特征值中引入误差. RSS 的不一致性会导致生成密钥存在不一致的比特位. 为了降低误码率我们引入了不一致性去除这一步骤,旨在去除通信双方的不一致 RSS 值.

下面我们具体分析可能会在量化过程中产生错误 RSS 值的特点. 首先,我们以平均线(mean)将 RSS 划分为两个区间,在 mean 的上下两侧分别选取 q^+ 和 q^-,并以 q^+ 和 q^- 为参考界限,去掉 q^+ 和 q^- 之间的 RSS. Alice 和 Bob 进行通信留下双方共同保留的 RSS 值以去除这些存在微小偏移的点,减少误码率. 其次,A、B 两方的 RSS 可能存在大幅跳变,如 Alice 处的 RSS 值处于 q^+ 之上而 Bob 处对应的 RSS 值却跳变到 q^- 之下. 一般情况下仅有单个 RSS 会发生这种跳变,而连续几位大幅跳变的概率是很小的. 因此我们引入 m 矫正因子,使得只有在连续 m 位发生大幅度跳变的情况下,才有可能产生误码率.

由于在信息收集阶段 Alice 和 Bob 会收集大量的 RSS 信息,为了更准确地对 RSS 进行处理,我们首先以 b 长度对 RSS 进行分块(block). 其次,为了去掉微小偏移且使得 Alice 和 Bob 在不泄露信息的情况下进行协商,我们定义标签函数 $R(x)$,假设标签函数的输入为收集到的 RSS 数组 $X = \{x_1, x_2, \cdots, x_k\}$,$x_i \in \mathrm{Z}$,这 k 个 RSS 值分属于 t 个块中,每个块的长度为 b,则标签函数 $R(x)$ 为:

$$R(x_i) = \begin{cases} 0, & x_i > q^+_{\lfloor i/b \rfloor} \\ 1, & x_i < q^-_{\lfloor i/b \rfloor} \end{cases} \tag{1}$$

定义参考界限数组为 $Q+ = \{q_1^+, q_2^+, \cdots, q_t^+\}$ 和 $Q- =$

$\{q_1^-, q_2^-, \cdots, q_t^-\}$,其中,$q_i^+ = mean_i + \alpha * std\_derivation_i$,$q_i^- = mean_i - \alpha * std\_derivation_i$,$mean_i$为第 i 个 block 的 RSS 平均值,α 是波动因子且 $\alpha \in (0,1)$,$std\_derivation_i$ 是第 i 个 block 的 RSS 标准差. 根据参考界限,所有大于 q_i^+ 的 RSS 值都被标记为 0,所有小于 q_i^- 的 RSS 值都被标记为 1,而处于两个参考界限之间的值被全部丢弃以去除有微小偏移的 RSS 值带来的影响. 随后 Alice 和 Bob 在自己的 0、1 序列中检查,根据矫正因子 m 来矫正序列以消除大幅跳变,并进行通信留下双方共同保留的 RSS 值.

具体的不一致性去除步骤描述如下:

(1) Alice 根据标签函数生成 0、1 序列,遇到连续 m 位相同的 0 或 1,则将中间位的序号记录在 $L_{a\_to\_b}$ 数组中,生成 $L_{a\_to\_b} = \{l_1, l_2, \cdots, l_a\}$ 并发送给 Bob;

(2) Bob 根据标签函数,记下所有满足连续 m 位相同的中间位序号,与 $L_{b\_to\_a}$ 数组进行对比,并剔除不一致的值,生成数组并发送给 Alice;

(3) $L_{b\_to\_a}$ 数组中所有序号所对应的 RSS 值,即为有效的、Alice 和 Bob 共同保留的 RSS 值,是下一步矢量量化的输入.

3.2　N 维矢量量化

为了提高密钥生成速率,大多方案都采用增加量化区间的做法,这种做法势必会引起误码率的大幅升高. 本文提出的 N 维矢量量化方法,在不增加量化区间的基础上,可以有效地利用信道信息,提高了密钥生成速率,同时不会引起误码率的升高.

假设 N 维矢量量化的输入是经过不一致性去除后的有效 RSS 数组 $Y = \{y_1, y_2, \cdots, y_d\}$,$y_i \in Z$,$Y$ 数组中的 RSS 值为 $L_{b\_to\_a}$ 数组中每个序号所对应的 RSS 值,则 Y 数组经过标签函数处理,对应的标签数组为:

$$R = \{R(y_1), R(y_2), \cdots, R(y_d)\}, R(y_i) \in (0,1).$$

对于 N 维矢量量化,建立 N 维矢量:

$$< y_i, y_{(i+\Delta_1) \bmod d}, y_{(i+\Delta_1+\Delta_2) \bmod d}, \cdots, y_{(i+\Delta_1+\cdots+\Delta_{N-1}) \bmod d} > \quad (2)$$

其中,$\Delta = \{\Delta_1, \Delta_2, \cdots, \Delta_{N-1}\}$,是 N 维矢量中 N 个分量的分量间隔,Δ_j 是第 j 个分量和第 $j+1$ 个分量之间的序号间隔. 则对于每一个输入 y_i,N 维矢量量化的输出为:

$$Q_N(y_i) = R(y_i) R(y_{(i+\Delta_1) \bmod d}) \cdots R(y_{(i+\Delta_1+\cdots+\Delta_{N-1}) \bmod d}) \quad (3)$$

N 维矢量量化器 $Q_N(Y)$ 的作用是将不连续的 RSS 值转化为 0、1 比特流,当 y_i 作为矢量量化的输入时,量化器以参数数组 Δ 为依据,向后查找第 2 到 N 个分量,组成 N 维矢量,并根据这 N 个分量的序号,得出相应的标签值,最后将这 N 个分量的标签值缝合在一起得到 N 位的比特流输出,算法 1 详细描述了这一矢量量化的过程.

算法 1　矢量量化

输入　N:量化维数

$Y = \{y_1, y_2, \cdots, y_d\}$:不一致性去除之后的有效 RSS 测量值

$R = \{R(y_1), R(y_2), \cdots, R(y_d)\}$:由标签函数产生的值序列,其中,$R(y_i) \in \{0,1\}$

$\Delta = \{\Delta_1, \Delta_2, \cdots, \Delta_{N-1}\}$:分量间隔,其中 Δ_i 是第 i 个分量和第 $i+1$ 个分量之间的间隔

输出　$B = Q_N(Y)$:Alice 和 Bob 产生的比特流

```
1. Δ_t ← 0   //初始化变量
2. for i = 1 to d do
3.     Q_N(y_i) ← R(y_i)
4.     Δ_t ← 0
5.     for j = 1 to N − 1
6.         Δ_t ← Δ_t + Δ_j
7.         Q_N(y_i) ← strcat(Q_N(y_i), R(y_{(i+Δ_t) mod d}))
8.     end for
9. end for
10. B ← Q_N(y_1) Q_N(y_2) ··· Q_N(y_d)
```

对于 N 维矢量量化,每一位的 RSS 输入,都会有 N 位的比特输出,RSS 数量与生成比特的数量比例为 1:N. 且 N 维矢量量化可以看成是在 N 维空间内建立起 N 维坐标系,空间被坐标轴分割成 2^N 个空间,这些空间可以分别用 2^N 个不重复的长度为 N 的格雷码[17]来表示.

N 个 RSS 分量组成的矢量可能落在 2^N 个空间中的任何一个,这样就可以把这个分量映射到对应的长度为 N 的格雷码上,从而完成一位 RSS 到 N 位比特流的量化.

我们以二维矢量量化为例说明矢量量化的过程. 如图 2 所示,当输入为 y_i 时,以 Δ_1 为间隔向后找到第二个分量 $y_{(i+\Delta_1) \bmod d}$ 组成二维矢量 $< y_i, y_{(i+\Delta_1) \bmod d} >$,并根据这两个分量的序号找到不一致性去除时得到的标签输出值 $R(y_i)$ 和 $R(y_{(i+\Delta_1) \bmod d})$,这两位的标签输出值 $R(y_i)$ $R(y_{i+\Delta_1})$ 即为 y_i 的二维矢量量化输出,所以二维矢量量化器为:

$$Q(y_i) = \begin{cases} 00, y_i > q^+_{[i/b]}, y_{(i+\Delta_1) \bmod d} > q^-_{[i/b]}, \text{第一象限} \\ 10, y_i < q^+_{[i/b]}, y_{(i+\Delta_1) \bmod d} > q^-_{[i/b]}, \text{第二象限} \\ 11, y_i < q^+_{[i/b]}, y_{(i+\Delta_1) \bmod d} < q^-_{[i/b]}, \text{第三象限} \\ 01, y_i > q^+_{[i/b]}, y_{(i+\Delta_1) \bmod d} < q^-_{[i/b]}, \text{第四象限} \end{cases} \quad (4)$$

如图 2 所示,以 $Q+$ 和 $Q-$ 为参考界限,将参考界限之间的点都去除,二维量化中以 Δ 为间隔,取两个 RSS 组成二维矢量 $< x_1, x_2 >$,若第一个分量 x_1 大于 $Q+$ 且第二个分量 x_2 小于 $Q-$ 则矢量 $< x_1, x_2 >$ 映射到第四象限,x_1 的二维矢量量化结果为 01.

3.3　基于模糊提取器的随机性增强处理

矢量量化的过程中进行了比特的重用,这使得生成的各比特串之间存在相互联系,密钥的随机性比较

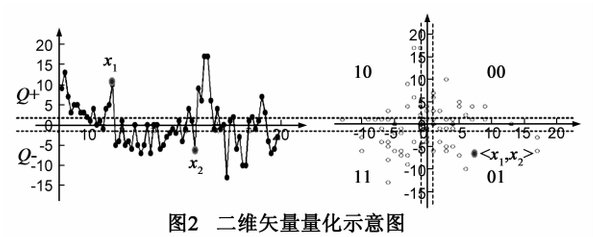


图2 二维矢量量化示意图

差. 我们在 N 维矢量量化之后引入了模糊提取器,在对密钥比特串进行纠错的同时也提高了其随机性. 模糊提取由一对过程 < Gen, Rep > 定义,它从输入w_0中以容错的方式可靠地提取出均匀分布的随机密钥 R. 为了从另一个与w_0充分接近的输入 w'中恢复出 R,模糊提取还同时输出公开信息 P,恢复过程必须有 P 的参与.

定义 1 一个参数为(M,l,t)的模糊提取由 Gen、Rep 两个过程组成:

(1)Gen 是一个概率生成过程,对于输入$w_0 \in M$,输出公开信息 P 和长度为 l 的随机秘密信息 R,即$(R,P) \leftarrow \mathrm{Gen}(w_0)$.

(2)Rep 是一个确定性恢复过程,对于输入公开信息 P 和与w_0充分接近的任意输入信息 w',输出对应的 R,即对所有的$w_0, w' \in M$,且满足 $d(w_0,w') \leqslant t$ 的 w',如果$(R,P) \leftarrow \mathrm{Gen}(w_0)$,就有 $R \leftarrow \mathrm{Rep}(w',P)$. 其中 $d(w_0,w') \leqslant t$ 表示w_0和 w'的距离不超过 t.

图 3 给出了标准模糊提取器的构造. 我们选取 BCH 码[18]来进行纠错,使用 SHA-1 函数对生成的比特流进行随机性增强,在 4.3 节中有详细描述.

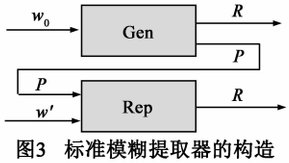


图3 标准模糊提取器的构造

我们从两个方面来考虑模糊提取器的安全问题:(1)输出值 R 是随机的,而且只在本地保存,可以保证其安全性;(2)攻击者不能根据 P 获得w_0值和 R 值,详细的安全分析可见文献[19].

4 实验及分析

在我们的实验中,Alice、Bob 和 Eve 三台主机的操作系统均为 ubuntu12.04,都使用相同型号的网卡 Atheros TL-WN650G,由 MadWifi 驱动,工作在 802.11g 模式下收发包. 我们对 MadWifi 进行了修改,添加了对 beacon 帧的回复,beacon_ack 携带与 beacon 相同的序列号,通过序列号的匹配,通信双方可以完成 RSS 的配对,beacon_ack 的长度为 49 字节. 实验在约有 20 台 PC 的实验室内进行,Bob 在 Alice 周围来回移动并以 100ms 为间隔向 Alice 发送 beacon 帧,Alice 收到 beacon 帧后会记录下序列号和对应的 RSS 值,并立刻给 Bob 发送一个 beacon_ack 的确认帧. Bob 收到确认帧后,做同样的操作. 整个测量过程中我们共收集到了 120000 个 RSS.

下面我们对方案的性能进行详细地分析,图 4 为最优情况下比特生成率随维数 N 变化的关系图,其中参数为 $b=80, \alpha=0.2, m=2, \Delta_1=\Delta_2=\cdots=\Delta_{N-1}=60$,矢量量化方案在在 7 维时可以达到 284% 的比特生成率,即收集到一个 RSS 信息可以生成 2.84bit 密钥. 由于 RSS 损失仅发生在不一致性去除步骤且矢量量化和模糊提取皆不会引起密钥长度损失,而量化步骤输入的每个有效 RSS 会对应 N 个比特的输出,所以当 b,α 和 m 固定时比特生成率与维数 N 为线性关系,如图 4 所示.

图 5 为比特误码率随维数 N 的变化关系图,可以看出,即使不经过纠错步骤,经过矢量量化这一步输出的比特流的误码率也仅为 0.116%,优于以往的所有方案. 由于矢量量化这一步产生的比特流误码率极低,所以才能在模糊提取这一步进行纠错,最终生成相同的密钥. 需要说明的是,经过不一致去除后,Alice 和 Bob 不一致的 RSS 数目固定,随着维数 N 的增大,比特流误码率保持不变.

4.1 不一致性去除分析

方案定义了标签函数,并通过 Alice 和 Bob 的通信来去掉可能引起不一致的 RSS 信息,在这个过程中主要引入了 b、α 和 m 三个参数.

首先,在实验中我们发现分块处理时引入的参数 b 并不是越小越好,分块过小或过大都会引起平均值计算不准确,从而使得误码率升高. 实验结果如图 6 所示,区间约为(15,100)时误码率都保持在较低水平.

其次,波动因子 α 越大,则参考界限 $Q+$ 和 $Q-$ 距离平均线的距离越远,存在微小偏差的点被剔除的概率就越高,误码率会下降,但同时两个参考界限之间的被舍弃的点也就越多,会造成比特生成率下降. 反之,α 越小意味着误码率增加和比特生成率升高,如图 7 和图 8 分别描述了 α 对比特生成率和误码率的影响.

最后,我们考虑矫正因子 m 的影响,m 越大则满足连续 m 位都大于 $Q+$ 或小于 $Q-$ 的点越少,协商生成的有效 RSS 值就越少,所以随着 m 增加比特生成率会大幅减小,而 m 增加意味着连续 m 位 RSS 大幅跳变的情况减少,所以误码率会相应下降;反之,m 减小意味着比特生成率增加和误码率减小. m 对比特生成率和误码率的影响情况如图 9 和图 10 所示.

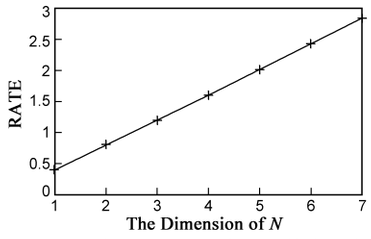


图4　N维矢量量化方案的比特生成率

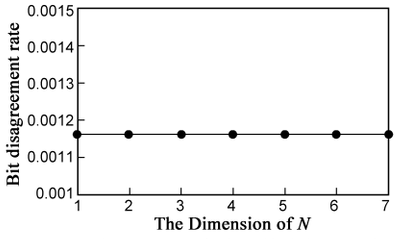


图5　N维矢量量化产生比特流的误码率

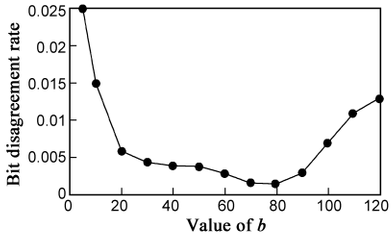


图6　b对误码率的影响

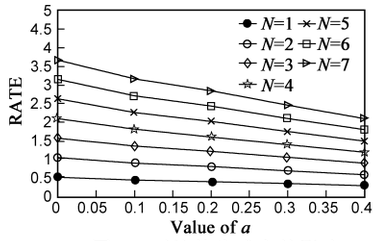


图7　α对比特生成率的影响

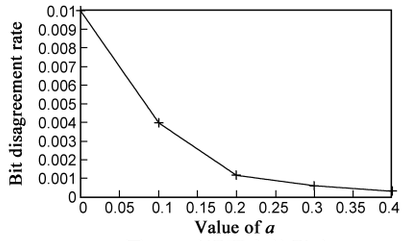


图8　α对误码率的影响

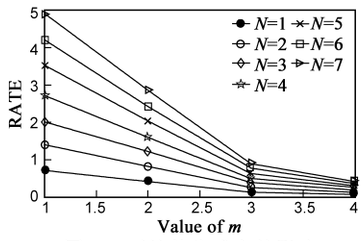


图9　m对比特生成率的影响

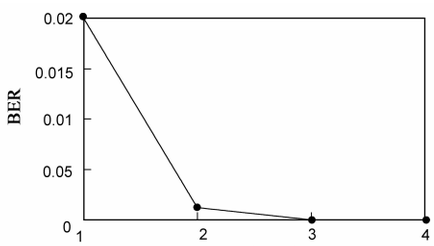


图10　m对误码率的影响

4.2　N 维矢量量化分析

N 维矢量量化意味着一位有效 RSS 的输入对应生成 N 个比特的输出,所以 N 越大方案的比特生成率就越高. 但为了确保所生成比特流的随机性,同已有方法一样,我们利用 NIST 的 9 项测试对不同维度下生成密钥的随机性进行了评估,对于每项,P-value值大于0. 01即表示通过测试. 结果如表 1 所示,当 $N \leqslant 7$ 的时候,所生成的比特流能够通过所有的检测项. 然而,当 $N \geqslant 8$ 的时候,每个比特被重用的次数太多,比特之间的独立性降低,随机性减弱,使得 NIST 测试项 FFT 值始终为 0,生成的比特流不能通过 NIST 随机性测试. 因此,从上述分析及测试结果可以看出,在我们的方案中 7 维是确保生成密钥随机性的一个上限.

N 维矢量中各个分量之间的间隔数组 $\Delta = \{\Delta_1, \Delta_2, \cdots, \Delta_{N-1}\}$ 对于生成的比特流也有影响,Δ_i不能小于相干时间,否则两个分量之间会存在关联,攻击者可能推测出其规律性,我们选取$\Delta_1 = \Delta_2 = \cdots = \Delta_{N-1} = 60$,此时任意两个 RSS 分量都处于相干时间之外,且每个 RSS 对应的矢量映射到 N 维空间坐标轴上的分布比较均匀,此外 Δ 的选取对比特生成率和误码率均没有影响.

表 1　不同维度生成密钥的 NIST 测试结果

P – value	一维	二维	三维	四维	五维	六维	七维	八维
Frequency	0. 48	0. 21	0. 35	0. 12	0. 68	0. 68	0. 31	0. 025
Block frequency	0. 99	0. 53	0. 58	0. 48	0. 48	0. 91	0. 87	0. 24
Cumulative sums(Fwd)	0. 39	0. 27	0. 14	0. 35	0. 18	0. 18	0. 68	0. 43
Cumulative sums(Rev)	0. 87	0. 58	0. 18	0. 74	0. 99	0. 16	0. 27	0. 16
Runs	0. 68	0. 91	0. 68	0. 74	0. 78	0. 24	0. 18	0. 53
Longest runs of ones	0. 96	0. 4	0. 87	0. 74	0. 27	0. 74	0. 96	0. 017
FFT	0. 28	0. 1	0. 23	0. 21	0. 31	0. 17	0. 07	0. 00
Approximate entropy	0. 53	0. 78	0. 12	0. 35	0. 91	0. 24	0. 35	0. 41
Serial	0. 21 ,0. 4	0. 83 ,0. 48	0. 63 ,0. 63	0. 78 ,0. 98	0. 21 ,0. 53	0. 87 ,0. 53	0. 03 ,0. 16	0. 04 ,0. 63

4.3 模糊提取器性能分析

本文的模糊提取器可描述为,Alice 通过矢量量化生成比特流w_0,作为 Gen 的输入,通过 BCH(23,12)纠错编码生成纠错序列P,并对w_0进行 SHA-1 哈希生成随机性强的密钥,哈希的输入与输出长度比为1: 1. 然后 Alice 将纠错序列 P 发送给 Bob. Bob 使用 P 对自己的量化生成比特流 w'进行纠错,恢复出w_0序列,然后同样进行 SHA-1 哈希生成相同的随机密钥. 在该过程中选取的 BCH(23,12)原码长度为 12,纠错码长度为 11,它的最小码距为 7,可以纠正 3 个错误. 我们将比特流分成每 12 位一组,生成对应的 11 位纠错码发送给通信的对方,这 11 位纠错码可以纠正误码率为 25% 的比特流,而本方案的误码率为 0.0116%,所以经过 BCH 码纠错后,Alice 和 Bob 生成的密钥完全一致,达到零误码率.

4.4 方案比较

为了能更好地体现本文方案的优势,我们选取了 Mathur[7] 的方案、Jana[8] 的方案 ASBG、Patwari 的方案 HRUBE[9] 和 ARUBE[10] 以及 Liu[14] 的方案等当前最典型的五个方案与我们的方案作对比研究.

我们从比特生成率、密钥误码率、密钥随机性三个方面来衡量方案的优劣,对比结果呈现在图 11 中. 经过对比,可以得出以下结论:

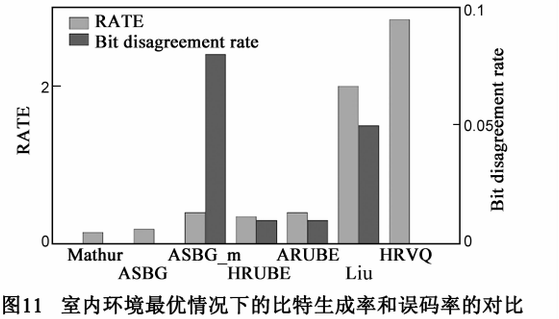


图11 室内环境最优情况下的比特生成率和误码率的对比

(1) Mathur 的方案误码率极低且生成的密钥随机,但是比特生成率仅有 15%.

(2) ASBG 的单比特提取方案性能与 Mathur 的方案相近,其多比特提取方案密钥速率虽然上升了,但误码率也大幅提高,当比特生成率达到 40% 时,误码率升高至 8%.

(3) Patwari 的 HRUBE 方案和 ARUBE 方案在误码率方面比前两个方案有了明显的提高,但这两个方案引入了额外的硬件 TelosB,且在 KLT 变换时奇异值分解的时间复杂度很高,可以达到 $O(n^3)$.

(4) Liu 的方案有着很高的比特生成率,接近 200%. 他的方案中指出 75% 的 RSS 值都处于上升或下降趋势中,但是在我们的实验中,处于上升或下降趋势的 RSS 比率随时间和周围环境的改变而不同,所以该方案的稳定性较差.

可以看出,我们的方案是唯一一个实现零误码率的方案,同时,将误码率限制在 1% 以内时,我们的方案达到了目前为止最高的比特生成率 284%,大大超出了已知的所有方案.

方案的可行性方面,从算法 1 可以看出,N 维矢量量化总的循环次数为 $d \times N$,其中 d 为经过不一致性去除后有效 RSS 数组的维数,而原始采集到 RSS 的个数 $Num = T \times f_b$(T 为 RSS 数据收集的时间,f_b 为 WiFi 中 beacon 帧发送的频率),因此 $d < T \times f_b$. 而 N 为矢量量化的维数,从上面分析可知,该值的上限为 7,即:$N \leqslant 7$. 因此 N 维矢量量化总的循环次数 $d \times N < 7 \times T \times f_b$,故 N 维矢量量化的时间复杂度 $O(d \times N) < O(7 \times T \times f_b)$ 为线性时间. 并且由于数组 R 共有 d 个成员,且 $d < T \times f_b$,所以该算法所需的空间复杂度为存储数组 R 所需的空间,即 $S(R) < S(T \times f_b)$. 因此该算法的空间复杂度也是线性的. 综上所述,矢量量化方法在现实中是可行的.

在实际应用场景中,单从密钥生成时间来看,本方案效率不如 EAP-TLS[20] 等传统协议方法. 但是,本方案可以很好地解决引言中所述的传统方案在应用中的问题,同时也能实现传统方案所不具备的匿名认证[21].

5 总结

针对物理层密钥生成方法中密钥生成速率与误码率之间的矛盾,本文提出了基于矢量量化的物理层密钥提取方案,通过不一致性去除减少通信双方不一致的信道特征值,利用矢量量化将信道信息转化为 0、1 比特流,并通过模糊提取器进行纠错及随机性增强处理. 实验表明相对于目前已知的最好方案,本方案将密钥生成速率提高了 6 倍,达到了 284% 的比特生成率,同时实现了零误码率,而且生成的密钥通过了 NIST 的 9 个随机性测试.

参考文献

[1] 孙宏,杨义先. 无线局域网协议 802.11 安全性分析[J]. 电子学报,2003,31(7):1098 - 1100.
Sun Hong, Yang Yixian. On the security of wireless network protocol 802.11[J]. Acta Electronica Sinica, 2003, 31(7):1098 - 1100. (in Chinese)

[2] Sayeed A, Perrig A. Secure wireless communications: Secret keys through multipath[A]. IEEE ICASSP[C]. Piscataway, NJ, USA: IEEE Press, 2008. 3013 - 3016.

[3] Shannon C E. Communication theory of secrecy systems [J]. J Bell Syst Tech, 1949, 28(4): 656 - 715.

[4] Xiao L, Greenstein L, Mandayam N, et al. A physical-layer

technique to enhance authentication for mobile terminals [A]. IEEE International Conference on Communications [C]. Piscataway, NJ, USA: IEEE, 2008. 1520 – 1524.

[5] Aono T, Higuchi K, Ohira T, et al. Wireless secret key generation exploiting reactance-domain scalar response of multipath fading channels[J]. IEEE Transactions on Antennas and Propagation, 2005, 53(11): 3776 – 3784.

[6] Hershey J, Hassan A, Yarlagadda R. Unconventional cryptographic keying variable management[J]. IEEE Transactions on Communications, 1995, 43(1): 3 – 6.

[7] Mathur S, Trappe W, Mandayam N, et al. Radio-telepathy: Extracting a secret key from an unauthenticated wireless channel[A]. ACM MobiCom[C]. New York, NY, USA: ACM Press, 2008. 128 – 139.

[8] Jana S, Pre S, Clark M, et al. On the effectiveness of secret key extraction from wireless signal strength in real environments[A]. ACM Mobi Com[C]. New York, NY, USA: ACM, 2009. 321 – 332.

[9] Patwari N, Croft J, Jana S, et al. High-rate uncorrelated bit extraction for shared secret key generation from channel measurements[J]. IEEE Transactions on Mobile Computing, 2010, 9(1): 17 – 30.

[10] Croft J, Patwari N, Kasera S. Robust uncorrelated bit extraction methodologies for wireless sensors[A]. ACM/IEEE ICNP[C]. New York, NY, USA: ACM Press, 2010. 70 – 81.

[11] Zhang J, Firooz M H, Patwari N, et al. Advancing wireless link signatures for location distinction[A]. ACM MobiCom[C]. New York, NY, USA: ACM, 2008. 26 – 37.

[12] Azimi-Sadjadi B, Kiayias A, Mercado A, et al. Robust key generation from signal envelopes in wireless networks [A]. ACM CCS[C]. New York, NY, USA: ACM, 2007. 401 – 410.

[13] Tope M, Mceachen J. Unconditionally secure communications over fading channels[A]. IEEE MILCOM[C]. Piscataway, NJ, USA: IEEE, 2001. 54 – 58.

[14] Liu H, Yang J, Wang Y, et al. Collaborative secret key extraction leveraging received signal strength in mobile wireless networks[A]. IEEE INFOCOM[C]. Piscataway, NJ, USA: IEEE, 2012. 927 – 935.

[15] Liu H, Yang J, Wang Y, et al. Group secret key generation via received signal strength: Protocols, achievable rates, and implementation[J]. IEEE Transactions on Mobile Computing, 2014, 13(12): 2820 – 2835.

[16] Dou L, Wei Y, Ni J. Multi-user wireless channel probing for shared key generation with a fuzzy controller[J]. Computer Networks, 2014, 58: 112 – 126.

[17] Ye C, Reznik A, Shah Y. Extracting secrecy from jointly Gaussian random variables[A]. IEEE ISIT[C]. Piscataway, NJ, USA: IEEE, 2006. 2593 – 2597.

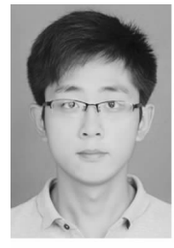
[18] Dodis Y, Ostrovsky R, Reyzin L, et al. Fuzzy extractors: How to generate strong keys from biometrics and other noisy data[J]. SIAM Journal on Computing, 2008, 38(1): 97 – 139.

[19] Dodis Y, Reyzin L, Smith A. Fuzzy extractors: How to generate strong keys from biometrics and other noisy data [A]. International Conference on the Theory and Applications of Cryptographic Techniques[C]. Heidelberg, GER: Springer, 2004. 523 – 540.

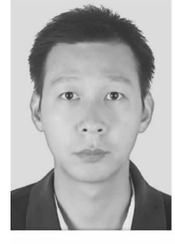
[20] Li X, Bao F, Li S, Ma J. FLAP: An efficient WLAN initial access authentication protocol[J]. IEEE Transactions on Parallel and Distributed Systems, 2014, 25(2): 488 – 497.

[21] 李兴华,尚昭辉,杨丹,马建峰. 利用无线物理层密钥增强 802.11i 的安全性[J]. 江苏大学学报(自然科学版), 2013, 34(4): 416 – 421.
Li Xinghua, Shang Zhaohui, Yang Dan, Ma Jianfeng. Security enhancement of 802.11i by wireless physical layer key[J]. Journal of Jiangsu University(Natural Science Edition), 2013, 34(4): 416 – 421. (in Chinese)

作者简介



李　鑫　男,1989 年生于山东省潍坊市,在读硕士研究生,研究方向为无线物理层安全、安卓安全.
E-mail: xdulixin@163.com



李兴华　男,1978 年生于河南省南阳市,博士,教授,博士生导师. 研究方向包括网络与信息安全、隐私保护、云计算等.
E-mail: xhli1@mail.xidian.edu.cn

杨　丹　女,1989 年生于陕西省咸阳市,硕士. 研究方向为无线物理层安全.
E-mail: danyang1989@gmail.com

马建峰　男,1963 年生于陕西省西安市,博士,教授,博士生导师. 研究方向包括信息安全、编码理论、密码学等.
E-mail: jfma@mail.xidian.edu.cn